



STIC Search Report

Biotech-Chem Library

File Copy
10/042059
updated

STIC Database Tracking Number: 148695

TO: David Lamberston
Location: rem/2b79/2c70
Art Unit: 1636
Friday, April 01, 2005

Case Serial Number: 10/042059

From: Barb O'Bryen
Location: Biotech-Chem Library
Remsen 1a69
Phone: 571-272-2518

barbara.obryen@uspto.gov

Search Notes

STIC-Biotech/ChemLib

148 695

From: Lambertson, David
Sent: Wednesday, March 23, 2005 1:32 PM
To: STIC-Biotech/ChemLib
Cc: Lambertson, David
Subject: Search Request

STIC
MAR 23 2005
1:32 PM

Search Request

Examiner's Name:	David Lambertson
Examiner #:	79514
Art Unit:	1636
Room #:	Remsen 02B79
Mailbox room#:	Remsen 02C70
Phone:	(571) 272-0771
Results Format:	paper

Serial # 10/042059

Please Search:

Nucleic Acid and Protein databases for:

SEQ ID NO: 1 and SEQ ID NO: 2

Including:

1. Default Search.
2. Interference Search.

Thanks,
Dave.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: March 31, 2005, 03:00:36 / Search time 204 Seconds
(without alignments)
3196.461 Million cell updates/sec

Title: US-10-042-059b-1

Perfect score: 1504

Sequence: 1 atgagctatcgaagcagca.....ctctctcgtcgttactag 843

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2105692 seqs, 366760381 residues

Total number of hits satisfying chosen parameters: 4211384

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame_n2p:model -DEV=xlh
-O=/cgn2.1/USPFO/US10042055/runat.30032005.103351.7699/app_query.fasta.1.1031
-DB=A Geneseq.16Dec04 -QFMT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOCL=0
-LOOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10042055_@CGN1.1.137_@runat.30032005.103351.7699 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOC
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAEXT=0.5 -DELOP=6 -DELEXT=7

Database : A_Geneseq.16Dec04:*

1: geneseqp1980s:*\n2: geneseqp1980s:*\n3: geneseqp2000s:*\n4: geneseqp2001s:*\n5: geneseqp2002s:*\n6: geneseqp2003as:*\n7: geneseqp2003bs:*\n8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1434	95.3	280	3	AA011451
2	921	61.2	268	8	ADP98857
3	587.5	39.1	266	8	ADP98857 C. albica
4	581.5	38.7	263	6	AA019566
5	508.5	33.8	257	8	ABU08093
6	508.5	33.8	257	8	ADJ55264
7	481.5	32.0	316	5	ABU91515
8	481.5	32.0	331	6	ABU08092
9	481.5	32.0	331	8	ADJ55262
10	466	31.0	334	5	AB092474

11	464	30.9	334	6	ABU08097
12	464	30.9	334	6	ADJ55271
13	423	28.1	261	8	ABU08094
14	423	28.1	261	8	ADJ55266
15	399	26.5	265	5	AA049654
16	399	26.5	265	5	AB093424
17	399	26.5	265	6	ABU08098
18	399	26.5	265	6	ADJ55272
19	399	26.5	330	5	AA049655
20	391	26.0	319	3	AA060603
21	391	26.0	325	3	AA060602
22	391	26.0	347	3	AA060601
23	313	20.8	207	6	ABU08095
24	313	20.8	207	6	ADJ55268
25	221	14.7	154	6	ABU08090
26	221	14.7	154	6	ADJ55258
27	197.5	13.1	132	6	ABU08089
28	197.5	13.1	132	6	ADJ55256
29	148	9.8	1938	6	AB098398
30	132	8.8	93	6	ABU08087
31	132	8.8	93	8	ADJ55252
32	127	8.4	248	7	AB079314
33	122	8.1	62	6	ABU08088
34	122	8.1	62	6	ADJ55254
35	120	8.0	1938	6	ABP76678
36	119	7.9	201	4	AAU51115
37	119	7.9	201	6	AB047634
38	118	8.0	1938	6	ABP76680
39	116.5	7.7	373	2	AAV07202
40	116	7.7	1938	6	ABP76679
41	115.5	7.7	1093	2	AA041001
42	114.5	7.6	446	2	AAV20059
43	114.5	7.6	469	2	AAV20058
44	113.5	7.5	539	7	AB078971
45	113	7.5	409	7	AB074668

ALIGNMENTS

RESULT 1	AA011451	standard; protein; 280 AA.
AC	AA011451;	
XX		
DT	12-SEP-2003 (revised)	
DT	01-MAR-2001 (first entry)	
XX		
DE	H. polymorpha chorismate mutase protein.	
XX		
KW	Chorismate mutase; prephenate; selection marker; auxotrophic yeast.	
XX		
OS	Pichia angusta.	
XX		
PN	WO200065071-A1.	
PD	02-NOV-2000.	
XX		
PF	27-APR-2000; 2000WO-EP003844.	
XX		
PR	27-APR-1999; 99DE-01019124.	
XX		
PA	(RHEI-) RHEIN BIOTECH GES NEUE BIOTECHNOLOGISCHE.	
XX		
PI	Gellissen G, Braus G, Pries R, Krapmann S, Strasser AW;	
XX		
DR	WPI. 2000-687355/67.	
XX		
DR	N-ESDB; AAC81949.	
XX		
PT	Nucleic acids encoding chorismate mutase, useful for preparing an	
PT	auxotrophic selection system for recombinant yeast and for recombinant	
PT	protein expression.	
XX		

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_nbp model

Run on: March 31, 2005, 04:36:43 ; Search time 99.5 Seconds
(without alignments)
1264.908 Million cell updates/sec

Title: US-10-042-059B-1
Perfect score: 1504
Sequence: 1 atgaccttaagaacccaga.....ctccctctgtctactag 843

Scoring table:
BLOSUM62
Xgapop 10.0, Ygapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 513545 segs, 74649064 residues

Total number of hits satisfying chosen parameters: 1027090

Minimum DB seg length: 0
Maximum DB seg length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+2bp.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10042059/runat_30032005_103354_7738/app_query.fasta_1.1031
-DB=Issued Patents_AA -QPMT=fastran -SUFFIX=real -MINMATCH=0.1 -LOOPCL=0
-LISTEXT=0 -UNITS=bites -START=1 -END=1 -MATRIX=bloum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTPMT=pct -NORML=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10042059_QCEN_1.1.30@runat_30032005_103354_7738 -NCP=6 -ICU=3
-NO MMAP -LARGOQUERY -NEG SCORES=0 -WAIT -DSEBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents AA:*
1: /cgn2_6/ptodata/1/aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/aa/PCTUS.COMB.pep:*
6: /cgn2_6/ptodata/1/aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	921	61.2	315	US-09-248-796A-18213	Sequence 18213, A
2	508.5	33.8	257	US-09-454-279-14	Sequence 14, Appl
3	481.5	32.0	331	US-09-454-279-12	Sequence 12, Appl
4	464	30.9	334	US-09-454-279-21	Sequence 21, Appl
5	423	28.1	261	US-09-454-279-16	Sequence 16, Appl
6	399	26.5	265	US-09-454-279-22	Sequence 22, Appl
7	313	20.8	207	US-09-454-279-18	Sequence 18, Appl
8	221	14.7	154	US-09-454-279-8	Sequence 8, Appl
9	197.5	13.1	132	US-09-454-279-6	Sequence 6, Appl
10	132	8.8	93	US-09-454-279-2	Sequence 2, Appl
11	127	8.4	248	US-09-252-991A-28060	Sequence 28060, A
12	122	8.1	62	US-09-454-279-4	Sequence 4, Appl

13	113.5	7.5	539	US-09-252-991A-27717	Sequence 27717, A
14	113	7.5	409	US-09-252-991A-23414	Sequence 23414, A
15	111	7.4	273	US-09-252-991A-22218	Sequence 22218, A
16	111	7.4	373	US-09-252-991A-28703	Sequence 28703, A
17	109.5	7.3	335	US-09-252-991A-23948	Sequence 23948, A
18	109.5	7.3	423	US-09-252-991A-23967	Sequence 23967, A
19	109	7.2	163	US-09-252-991A-24968	Sequence 24968, A
20	108	7.2	1480	US-09-191-647-7	Sequence 7, Appl
21	108	7.2	1480	US-09-540-245A-7	Sequence 7, Appl
22	108	7.2	1480	US-09-540-153-7	Sequence 7, Appl
23	108	7.2	1480	US-09-182-024A-5	Sequence 5, Appl
24	108	7.2	1480	PCT-US91-09005-2	Sequence 2, Appl
25	107.5	7.1	296	US-09-252-991A-32162	Sequence 32162, A
26	107.5	7.1	2150	US-09-321-987B-2	Sequence 155, App
27	107.5	7.1	2165	US-09-800-729-155	Sequence 24278, A
28	107	7.1	375	US-09-252-991A-24278	Sequence 21106, A
29	107	7.1	697	US-09-252-991A-21106	Sequence 26790, A
30	106.5	7.1	369	US-09-252-991A-20790	Sequence 20790, A
31	106	7.0	589	US-09-252-991A-18210	Sequence 18210, A
32	106	7.0	659	US-09-252-991A-25283	Sequence 25283, A
33	105.5	7.0	383	US-09-252-991A-29621	Sequence 29621, A
34	105.5	7.0	565	US-09-252-991A-17762	Sequence 17762, A
35	105	7.0	583	US-09-252-991A-28056	Sequence 28056, A
36	105	7.0	661	US-09-252-991A-18225	Sequence 18225, A
37	104.5	6.9	343	US-09-252-991A-17876	Sequence 17876, A
38	104.5	6.9	532	US-09-252-991A-28042	Sequence 28042, A
39	104	6.9	205	US-09-252-991A-28770	Sequence 28770, A
40	104	6.9	208	US-09-252-991A-25609	Sequence 25609, A
41	104	6.9	1523	US-09-538-092-955	Sequence 955, App
42	103.5	6.9	296	US-09-252-991A-21951	Sequence 21951, A
43	103.5	6.9	296	US-09-252-991A-17878	Sequence 17878, A
44	103.5	6.9	437	US-09-252-991A-32814	Sequence 32814, A
45	103	6.8	266	US-09-252-991A-19128	Sequence 19128, A

ALIGNMENTS

RESULT 1
US-09-248-796A-18213
Sequence 18213, Application US/09248796A
Patent No. 6747137
GENERAL INFORMATION:
APPLICANT: Kelch Weinstock et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICA
FILE REFERENCE: 107196.132
CURRENT FILING DATE: US/09/248, 796A
PRIOR FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: US 60/074, 725
PRIOR FILING DATE: 1998-02-13
PRIOR APPLICATION NUMBER: US 60/096, 409
PRIOR FILING DATE: 1998-08-13
NUMBER OF SEQ ID NOS: 28208
SEQ ID NO 18213
LENGTH: 315
TYPE: PRT
ORGANISM: Candida albicans
US-09-248-796A-18213

Alignment Scores:

Pred. No.: 4,76e-89
Score: 921.00
Percent Similarity: 82.63%
Best Local Similarity: 65.64%
Query Match: 61.24%
DB: 4
Gaps: 0

US-10-042-059B-1 (1-843) x US-09-248-796A-18213 (1-315)

QY 1 ATGACCTTATGAGCCGAGAAAGTGTGACCTTGCAACATTAGATGCTTGTC 60
DB 48 MetaspPmetLysPProGluThrValIeuaspPleuAlaSerIleArgGlnAlaIeuVal 67

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: March 31, 2005, 05:46:38 ; Search time 275 Seconds
(without alignments)
2032.966 Million cell updates/sec

Title: US-10-042-059B-1

Perfect score: 1504
Sequence: 1 atgagcttcatgaagcaga.....tctctctgtgtactag 843

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Rgapop 6.0, Rgapext 7.0
Delop 6.0, Delext 7.0

Searched: 1413372 seqs, 331592847 residues

Total number of hits satisfying chosen parameters: 2826744

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xli
-O=/cgn2.1/USPRO.spb00/US10042059/runat 30032005.103356.7783/arp.query.fasta_1.1031
-DB=Published Applications AA -QMT=faefan -SUFFIX=rapb -MINMATCH=0.1
-TRANSL=0 -LOOPEXT=0 -UNITS=DIC -START=1 -END=1 -MATRIX=DIOSUM62
-TRANS=numan40.cdi -LIST=45 -DOCLIGN=200 -THR SCORE=ptc -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10042059 @CGN 1.1.103 @runat 30032005.103356.7783
-NCPU=6 -ICPU=3 -NO-WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOC=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELXT=7

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/CTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result Query

No.	Score	Match	Length	DB	ID	Description
1	1434	95.3	280	13	US-10-042-059A-2	Sequence 2, Appl1
2	921	61.2	268	17	US-10-741-849-7032	Sequence 7032, Ap
3	508.5	33.8	257	9	US-09-454-279-14	Sequence 14, Appl
4	508.5	33.8	257	15	US-10-624-061-14	Sequence 14, Appl
5	508.5	33.8	313	16	US-10-437-963-196439	Sequence 196439, A
6	497.5	33.1	320	15	US-10-425-114-68586	Sequence 68586, A
7	495.5	32.9	315	15	US-10-424-599-266662	Sequence 266662, A
8	486	32.3	350	16	US-10-767-701-45917	Sequence 45917, A
9	481.5	32.0	331	15	US-09-454-279-12	Sequence 12, Appl
10	481.5	32.0	331	15	US-10-624-061-12	Sequence 12, Appl
11	464	30.9	334	9	US-09-454-279-21	Sequence 21, Appl
12	464	30.9	334	9	US-10-624-061-21	Sequence 21, Appl
13	423	28.1	261	9	US-09-454-279-16	Sequence 16, Appl
14	423	28.1	261	15	US-10-624-061-16	Sequence 16, Appl
15	423	28.1	297	15	US-10-425-114-48703	Sequence 48703, A
16	419	27.9	261	15	US-10-424-599-235878	Sequence 235878, A
17	408.5	27.2	255	16	US-10-437-963-151855	Sequence 151855, A
18	399	26.5	265	9	US-09-454-279-22	Sequence 22, Appl
19	399	26.5	265	15	US-10-624-061-22	Sequence 22, Appl
20	374.5	24.9	313	15	US-10-425-114-65608	Sequence 65608, A
21	373	24.8	264	15	US-10-425-114-65608	Sequence 65608, A
22	356	23.7	306	16	US-10-437-963-186047	Sequence 186047, A
23	348	23.1	316	16	US-10-437-963-186040	Sequence 186040, A
24	344.5	22.9	186	15	US-10-425-114-52677	Sequence 52677, A
25	328	21.8	199	15	US-10-425-114-52677	Sequence 43606, A
26	313	20.8	207	9	US-09-454-279-18	Sequence 18, Appl
27	313	20.8	207	15	US-10-624-061-18	Sequence 18, Appl
28	297.5	19.8	489	16	US-10-437-963-103185	Sequence 103185, A
29	282	18.8	210	16	US-10-767-701-32934	Sequence 32934, A
30	221	14.7	154	9	US-09-454-279-8	Sequence 8, Appl1
31	221	14.7	154	15	US-10-624-061-8	Sequence 8, Appl1
32	197.5	13.1	132	9	US-09-454-279-6	Sequence 6, Appl1
33	197.5	13.1	132	15	US-10-624-061-6	Sequence 6, Appl1
34	178	11.8	109	15	US-10-424-599-224668	Sequence 224668, A
35	157.5	10.5	97	15	US-10-425-114-38176	Sequence 38176, A
36	151.5	10.1	19695	15	US-10-084-846A-3	Sequence 3, Appl1
37	132	8.8	93	9	US-09-454-279-2	Sequence 2, Appl1
38	132	8.8	93	15	US-10-624-061-2	Sequence 2, Appl1
39	130.5	8.9	278	15	US-10-425-114-48480	Sequence 48480, A
40	130.5	8.9	279	15	US-10-425-114-43408	Sequence 43408, A
41	124.5	8.3	797	14	US-10-156-761-10907	Sequence 10907, A
42	122	8.1	62	9	US-09-454-279-4	Sequence 4, Appl1
43	122	8.1	62	15	US-10-624-061-4	Sequence 4, Appl1
44	119	7.9	182	16	US-10-437-963-116837	Sequence 116837, A
45	117.5	7.8	459	16	US-10-437-963-153917	Sequence 153917, A

ALIGNMENTS

RESULT 1
US-10-042-059A-2
Sequence 2, Application US/10042059A
Publication No. US20020197704A1
GENERAL INFORMATION:
APPLICANT: Rhein BioTech Gesellschaft fur neue biotechnologische Prozesse und
APPLICANT: Produkt mbH
TITLE OF INVENTION: Nucleic Acid Molecule Containing a Nucleic Acid Coding for a Pol
FILE REFERENCE: PA30558US-019
CURRENT APPLICATION NUMBER: US/10/042, 059A
CURRENT FILING DATE: 2001-10-25
PRIOR APPLICATION NUMBER: DE 199 19 124.7
PRIOR FILING DATE: 1999-04-27
NUMBER OF SEQ ID NOS: 3
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 280
TYPE: PRT
ORGANISM: Hansenula polymorpha
US-10-042-059A-2

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: March 31, 2005, 04:14:20 ; Search time 44 Seconds
(without alignments)
3686.851 Million cell updates/sec

Title: US-10-042-059B-1

Perfect score: 1504
Sequence: 1 atgagacttcatgagacaga.....ctccctcgtgctgtactag 843

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283416 segs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seg length: 0
Maximum DB seg length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p: model -DSV=xlh
-O=/cgn2.1/USPTO.epool/US10042059/runat_30032005_103353_7727/app_query.fasta_1.1031
-DB=PIR_9 -OPMT=fastan -SUFFIX=tpir -MINMATCH=0.1 -LOOPCH=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pct -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10042059 @CCN 1.1 25 @runat_30032005_103353_7727 -NCPU=6 -ICPU=3
-NO MMAP -LARGEBUFFER -NEG SCORES=0 -WAIT -DSFBLOCK=100 -LONGIGLO
-DSV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCGAPOP=6
-FCGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
1: PIR_79:*
2: p1r1:*
3: p1r2:*
4: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	776.5	51.6	256	2 A45921	chorismate mutase
2	558.5	37.1	251	2 T37784	chorismate mutase
3	464	30.9	334	2 S38958	chorismate mutase
4	399	26.5	265	2 T50796	chorismate mutase
5	166	11.0	82	2 T14902	chorismate mutase
6	119	7.9	70	2 T14901	hypothetical prote
7	114.5	7.6	469	2 C70109	hypothetical prote
8	108	7.2	1480	2 A3665	hypothetical prote
9	107.5	7.1	2165	2 T21371	slit protein 1 pre
10	104	6.9	313	2 B75138	hypothetical prote
11	104	6.9	336	2 S04739	tetrahydropteroyl
12	104	6.9	1203	2 S26650	site-specific DNA-
13	103.5	6.9	1732	2 T14039	DNA-binding protei
14	103	6.8	560	1 P69059	protein kinase (EC arginine-CRNA ligase

15	102.5	6.8	376	2 F75503	conserved hypochet
16	102	6.8	1033	2 T37715	actin-interacting
17	101.5	6.7	625	2 T32739	hypothetical prote
18	101.5	6.7	728	2 S26427	70K structural prote
19	99.5	6.8	534	2 T39903	serine-rich protei
20	99	6.6	2142	1 ZLVNVP	genome polyprotein
21	98.5	6.5	424	2 H87520	hypothetical prote
22	98	6.5	1702	2 T14050	hypothetical prote
23	98	6.5	1791	2 T02345	protein kinase (EC
24	97.5	6.5	415	2 D95275	hypothetical prote
25	97.5	6.5	492	2 A87471	probable two-compo
26	97.5	6.5	1286	2 T37366	hypothetical prote
27	97	6.4	323	2 D90527	RNA polymerase sub
28	97	6.4	514	1 W2ML5	glycerol-3-phospha
29	97	6.4	1167	2 E96963	E2 protein - human
30	96.5	6.4	829	2 S72366	DNA topoisomerase III
31	96.5	6.4	932	2 F84465	hypothetical prote
32	96.5	6.6	2232	2 T34434	hypothetical prote
33	96	6.4	528	2 B75310	conserved hypochet
34	96	6.4	680	2 S21323	probable endoglucan
35	96	6.4	808	2 H64474	hypothetical prote
36	96	6.4	1286	1 H36845	DNA-directed RNA p
37	96	6.4	1286	2 T28521	DNA-directed RNA p
38	96	6.4	1286	2 A72161	M6R protein - vari
39	96	6.4	2127	1 ZLVNBP	genome polyprotein
40	95.5	6.3	280	2 H72593	hypothetical prote
41	95	6.3	623	2 S55652	transcription cont
42	94.5	6.3	310	2 S43865	cytochrome b, typ
43	94.5	6.4	402	2 E86185	hypothetical prote
44	94.5	6.3	790	2 T19683	hypothetical prote
45	94.5	6.3	1287	1 RMY247	DNA-directed RNA p

ALIGNMENTS

RESULT 1

A45921
Chorismate mutase (EC 5.4.99.5) - yeast (*Saccharomyces cerevisiae*)

N/Alternate names: protein YP9499.1sc; protein YPR060C

C/Species: *Saccharomyces cerevisiae*

C/Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #ext_change 09-Jul-2004

C/Accession: A45921; S54081

R/Schmidheini, T.; Sperisen, P.; Paravicini, G.; Huetter, R.; Braus, G.

J. Bacteriol. 171, 1245-1253, 1969

A/Title: A single point mutation results in a constitutively activated and feedback-ree

A/Reference number: A45921; MUID:89155418; PMID:2646272

A/Accession: A45921

A/Molecule type: DNA

A/Residues: 1-256 <SCH>

A/Cross-references: UNIPROT:P32178; GB:M24517; NID:g295576; PIDN:AAB59309.1; PID:g29557

R/Baddock, K.; Churcher, C.M.

submitted to the EMBL Data Library, May 1995

A/Reference number: S54059

A/Accession: S54081

A/Molecule type: DNA

A/Residues: 1-256 <BAD>

A/Cross-references: EMBL:249219; NID:g805025; PIDN:CAA89177.1; PID:g805040; MIPS:YPR06C

A/Experimental source: strain AB972

C/Genetics:

A/Gene: SGD:ARO7

A/Cross-references: SGD:S0006264; MIPS:YPR060C

A/Map position: 16R

C/Function:

A/Description: intramolecular transferase; isomerase

A/Pathway: aromatic amino acid biosynthesis

C/Superfamily: chorismate mutase of the AroQ class, eukaryotic type

C/Keywords: aromatic amino acid biosynthesis; intramolecular transferase; isomerase

Alignment Scores:

Pred. No.: 6.12e-55 Length: 256
Score: 776.50 Matches: 152
Percent Similarity: 74.13% Conservative: 40
Best Local Similarity: 58.69% Mismatches: 62

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 30, 2005, 21:05:55 ; Search time 4115 Seconds

(without alignments)
9926.548 Million cell updates/sec

Title: US-10-042-059B-1

Perfect score: 843

Sequence: 1 atggaacttatgagccagca.....tctccctcgtgctgactag 843

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing:

Minimum Match 0%

Listing first 45 summaries

Database : GenBank:*

1: gb_ba:*
2: gb_hgt:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_stb:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	843	100.0	843	6	BD266599 Nucleic a
2	843	100.0	843	6	AX043758 Sequence
3	843	100.0	1652	8	AF204738 Pichia an
4	843	100.0	1655	6	BD266600 Nucleic a
5	843	100.0	1655	6	AX043760 Sequence
6	325.2	38.6	110000	8	CR382131 20
7	281.8	33.4	110000	8	CR382138 20
8	270	32.0	948	6	AR548979
9	256.4	30.4	110000	8	AE016817_01
10	217.4	25.8	2059	8	YSCAR07A
11	217.4	25.8	43776	8	SC9499X
12	217.4	25.8	165536	8	SCCHXYVI
13	217.2	25.8	771	8	AY633179
14	210.4	25.0	110000	8	CR382122_07
15	189	22.4	110000	8	CR380957_11
16	156.4	18.6	869	8	BD178335
17	156.4	18.6	869	8	AB116236
18	141.8	16.8	1348	8	AK105512
19	131.4	15.6	1231	6	AR404687

20	131.4	15.6	1250	8	AK068983
21	113.6	13.5	3167	8	AF133241
22	111.2	12.3	1223	6	AR404686
23	110	13.0	1000	8	L47356
24	107.6	12.8	951	8	BT005306
25	107.6	12.8	1143	8	AK117860
26	104.4	12.4	1217	6	AR236635
27	104.4	12.4	1217	6	AR236641
28	104.4	12.4	1217	6	AX464576
29	104.4	12.4	1217	6	AX464582
30	104.4	12.4	1217	6	AF131219
31	100.6	11.9	110000	2	AP006499_05
32	99.8	11.8	1020	6	AR404688
33	97.4	11.6	798	6	AX506277
34	97.4	11.6	829	8	AY133840
35	97.4	11.6	993	6	AX343935
36	97.4	11.6	1006	6	AR236634
37	97.4	11.6	1006	6	AR236640
38	97.4	11.6	1006	6	AX343933
39	97.4	11.6	1006	6	AX464575
40	97.4	11.6	1006	6	AX464581
41	97.4	11.6	1006	8	L47355
42	97.4	11.6	1055	8	AY065238
43	93.6	11.1	1005	8	AX506042
44	93.6	11.1	1207	6	AR236633
45	93.6	11.1	1207	6	AR236639

ALIGNMENTS

RESULT 1
BD266599
LOCUS
DEFINITION
Nucleic acid molecule, containing a nucleic acid which codes for a polypeptide with chorismate mutase activity.

ACCESSION
BD266599.1 GI:33076367

VERSION
JP 2002542788-A/1.

SOURCE
Pichia angusta

ORGANISM
Pichia angusta

REFERENCE
Gellissen, G., Braus, G., Pries, R., Krappmann, S. and Strasser, A.W.

AUTHORS
Nucleic acid molecule, containing a nucleic acid which codes for a polypeptide with chorismate mutase activity

TITLE
Patent: JP 2002542788-A 1 17-DEC-2002.

JOURNAL
RHEIN BIOTECH GESELLSCHAFT FUER NEUR BIOTECHNOLOGISCHE PROZESSE UND

COMMENT
PRODUKT MEH

OS
PN JP 2002542788-A/1

PD
PD 17-DEC-2002

PP
PP 27-APR-2000 JP 2000614405

PR
PR 27-APR-1999 DE 199 19 124.7

PI
PI GERD GELIJSSEN, GERHARD BRAUS, RALPH PRIES, SVEN KRAPPMANN, PI

PC
PC C12N15/09, C07K16/40, C12N1/15, C12N1/19, C12N1/21, C12N5/10 PC

PC
PC C12N5/00

CC
CC Nucleic acid molecule, containing a nucleic acid which codes

CC
CC polypeptide with chorismate mutase activity

FT
FT source

FT
FT source

FEATURES
source

ORIGIN
1..843
/organism="Pichia angusta"
/mol_type="genomic DNA"
/db_xref="taxon:4905"

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 30, 2005, 18:17:40 ; Search time 588 Seconds
(without alignments)
8486.976 Million cell updates/sec

Title: US-10-042-059b-1

Perfect score: 843
Sequence: 1 atgagcttataagacgcaga.....tctctctgcgtctgactag 843

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

1: geneeqn19808:*\n2: geneeqn19908:*\n3: geneeqn20008:*\n4: geneeqn20018:*\n5: geneeqn20028:*\n6: geneeqn20038:*\n7: geneeqn20048:*\n8: geneeqn20058:*\n9: geneeqn20068:*\n10: geneeqn20078:*\n11: geneeqn20088:*\n12: geneeqn20098:*\n13: geneeqn20108:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	843	100.0	843	3	AAC81949
2	843	100.0	1655	3	AAC81950
3	270	32.0	807	12	ADP98547
4	183.8	21.7	801	13	ADR85510
5	156.4	18.6	869	8	AAH50202
6	141.2	16.7	989	13	ADR84923
7	141.2	16.7	2920	8	ABT17891
8	141.2	16.7	2987	8	ABT19705
9	141.2	16.7	6989	13	ADR84336
10	131.4	15.6	1231	10	ABX93057
11	131.4	15.6	1231	12	ADJ55263
12	116.8	13.9	658	3	AAH14469
13	116.8	13.9	1223	10	ABX93056
14	111.2	12.3	1223	12	ADJ55261
15	104.4	12.4	1217	6	ABA91395
16	104.4	12.4	1217	6	ABA91401
17	99.8	11.8	1020	10	ABX93058
18	99.8	11.8	1020	12	ADJ55265
19	97.4	11.6	798	6	ABT13167
20	97.4	11.6	993	6	ABA99614

21.	97.4	11.6	1006	6	ABA99613
22	97.4	11.6	1006	6	ABA91400
23	97.4	11.6	1006	6	ABA91394
24	93.6	11.1	1005	6	ABT12932
25	93.6	11.1	1207	6	ABA91393
26	93.6	11.1	1207	6	ABA91399
27	93.6	11.1	1274	3	AAC33324
28	73	8.7	580	13	ACN57510
29	69.8	8.3	622	13	ACN50132
30	64.2	7.6	600	13	ADR64815
31	61.8	7.3	258	6	ABT171526
32	61	7.2	780	10	ABX93059
33	61	7.2	780	12	ADJ55267
34	59.8	7.1	579	10	ABX93054
35	59.8	7.1	579	12	ADJ55257
36	56.2	6.7	584	13	ACN50205
37	54.8	6.5	542	13	ACN57601
38	53.4	6.3	525	10	ABX93053
39	53.4	6.3	525	12	ADJ55255
40	52.6	6.2	524	13	ADR64235
41	49.4	5.9	2000	8	ADA71938
42	48.8	5.8	864	12	ADJ44958
43	47.6	5.6	1379	6	ABA91794
44	47	5.6	5858	8	ADJ55726
45	46.8	5.6	180385	10	ADL13931

ALIGNMENTS

RESULT 1
AAC81949
ID AAC81949 standard; DNA, 843 BP.
XX
XX AAC81949;
AC
AC 15-SEP-2003 (revised)
DT 01-MAR-2001 (first entry)
XX
XX H. polymorpha chorismate mutase DNA.
XX
XX Chorismate mutase; prephenate; selection marker; auxotrophic yeast; ds.
XX
XX Pichia angusta.
OS
XX WO200065071-A1.
XX
XX 02-NOV-2000.
XX
XX 27-APR-2000; 2000WO-EP003844.
XX
XX 27-APR-1999; 99DE-01019124.
XX
XX (RHEI-) RHEIN BIOTECH GRS NEUE BIOTECHNOLOGISCHE.
XX
XX Gellissen G, Braus G, Pries R, Kraepmann S, Strasser AM;
XX WPI; 2000-687355/67.
XX P-PSDB; AAB11451.
XX
XX Nucleic acids encoding chorismate mutase, useful for preparing an
XX auxotrophic selection system for recombinant yeast and for recombinant
XX protein expression.
XX
XX Claim 1a; Page 57; 63pp; German.
XX
XX This invention describes novel nucleic acids (I) that encode a
XX polypeptide (II) with chorismate mutase (CM) activity (or its
XX complementary strand). CM catalyzes conversion of chorismate to
XX prephenate, an essential precursor for Phe and Tyr. (I) is a selection
XX marker for construction of corresponding auxotrophic yeast (requiring Phe
XX and Tyr) that are used for recombinant production of proteins. (I) allows
XX selection of transformed yeast on simple media. (Updated on 15-SEP-2003

Db 442 ATATTGGCCAATATCTGATGAATTAACTTAATCTGAATAATGAAGTTTATGTT 501

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 31, 2005, 02:00:01 ; Search time 3261 Seconds
(without alignments)
1564.628 Million cell updates/sec

Title: US-10-042-059B-1

Perfect score: 843
Sequence: 1 atggaacttcatgagccaga.....tctcctctgctgctactag 843

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5607317 seqs, 302624599 residues

Total number of hits satisfying chosen parameters: 11214634

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications NA:
1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/prodata/1/pubpna/US06_NEW_PUB.seq:*
3: /cgn2_6/prodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/prodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/prodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:*
7: /cgn2_6/prodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
12: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
13: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
14: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
15: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
16: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
17: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
18: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
19: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
20: /cgn2_6/prodata/1/pubpna/US11_NEW_PUB.seq:*
21: /cgn2_6/prodata/1/pubpna/US11_NEW_PUB.seq:*
22: /cgn2_6/prodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	843	100.0	843	US-10-042-059A-1	Sequence 1, Appl1
2	843	100.0	1655	US-10-042-059A-3	Sequence 3, Appl1
3	270	32.0	807	US-10-741-849-6032	Sequence 6032, Ap
4	156.4	18.6	869	US-10-472-587-26	Sequence 26, Appl
5	141.2	16.7	2920	US-10-128-714-249	Sequence 249, Appl
6	141.2	16.7	2987	US-10-128-714-249	Sequence 5249, Ap
7	131.4	15.6	1016	US-10-437-963-93956	Sequence 93956, A
8	131.4	15.6	1231	US-09-454-279-13	Sequence 13, Appl
9	131.4	15.6	1231	US-10-624-061-13	Sequence 13, Appl
10	118.2	14.0	1423	US-10-424-599-123820	Sequence 123820,
11	116.8	13.9	658	US-10-653-047-6992	Sequence 6992, Ap

12	116.6	13.8	1146	US-10-425-114-7285	Sequence 7285, Ap
13	116	13.8	1376	US-10-425-115-102347	Sequence 102347,
14	112.2	13.3	1548	US-10-739-930-4455	Sequence 4455, Ap
15	111.2	13.2	1223	US-09-454-279-11	Sequence 11, Appl
16	111.2	13.2	1223	US-10-624-061-11	Sequence 11, Appl
17	111.2	13.2	1500	US-10-767-701-14353	Sequence 14353, A
18	110.6	13.1	1010	US-10-425-115-102351	Sequence 102351,
19	109.6	13.0	1397	US-10-425-115-102346	Sequence 102346,
20	104.4	12.4	1217	US-10-267-763-4	Sequence 4, Appl1
21	104.4	12.4	1217	US-10-267-763-10	Sequence 10, Appl
22	103.2	12.2	1368	US-10-425-115-102349	Sequence 102349,
23	99.8	11.8	1020	US-09-454-279-15	Sequence 15, Appl
24	99.8	11.8	1020	US-10-425-114-7639	Sequence 7639, Ap
25	99.8	11.8	1079	US-10-425-114-7639	Sequence 7639, Ap
26	97.4	11.6	798	US-09-938-842A-972	Sequence 972, App
27	97.4	11.6	798	US-09-938-842A-972	Sequence 972, App
28	97.4	11.6	1006	US-10-267-763-3	Sequence 3, Appl1
29	97.4	11.6	1006	US-10-267-763-9	Sequence 9, Appl1
30	97	11.5	1430	US-10-739-930-4454	Sequence 4454, Ap
31	93.6	11.1	1005	US-09-938-842A-737	Sequence 737, App
32	93.6	11.1	1005	US-09-938-842A-737	Sequence 737, App
33	93.6	11.1	1207	US-10-267-763-2	Sequence 2, Appl1
34	93.6	11.1	1207	US-10-267-763-8	Sequence 8, Appl1
35	93.6	11.1	1299	US-10-739-930-621	Sequence 621, App
36	87.6	10.4	1653	US-10-424-599-93036	Sequence 93036, A
37	81	9.6	880	US-10-425-114-3898	Sequence 3898, Ap
38	80	9.5	1314	US-10-425-115-151470	Sequence 151470,
39	78.4	9.3	1197	US-10-425-114-27576	Sequence 27576, A
40	77.2	9.2	524	US-10-425-115-102342	Sequence 102342, A
41	76.2	9.0	768	US-10-437-963-48372	Sequence 48372, A
42	74.8	8.9	2154	US-10-437-963-83557	Sequence 83557, A
43	74.2	8.8	1066	US-10-425-114-1332	Sequence 1332, Ap
44	74.2	8.8	1083	US-10-425-115-115559	Sequence 115559,
45	73	8.7	580	US-10-021-323-12291	Sequence 12291, A

ALIGNMENTS

RESULT 1
US-10-042-059A-1
Sequence 1, Application US/10042059A
Publication No. US20020197704A1
GENERAL INFORMATION:
APPLICANT: Rhein Biotech Gesellschaft für neue biotechnologische Prozesse und
APPLICANT: Prokrite mH
TITLE OF INVENTION: Nucleic Acid Molecule Containing a Nucleic Acid Coding for a Pol
TITLE OF INVENTION: with Chorismante Mutase Activity
FILE REFERENCE: P4055805-019
CURRENT APPLICATION NUMBER: US/10/042,059A
PRIOR FILING DATE: 2001-10-25
PRIOR APPLICATION NUMBER: DE 199 19 124.7
NUMBER OF SEQ ID NOS: 3
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 1
LENGTH: 843
TYPE: DNA
ORGANISM: Hansenula polymorpha
US-10-042-059A-1

Query Match 100.0%; Score 843; DB 13; Length 843;
Best Local Similarity 100.0%; Pred.No. 6.4e-267;
Matches 843; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGACTTATGAAGCCAGAAACAGTGTGACCTTGGCAACTTGAATGCGCTTGGTC 60
DB 1 ATGGACTTATGAAGCCAGAAACAGTGTGACCTTGGCAACTTGAATGCGCTTGGTC 60
QY 61 CGGATGAGAGATACATCATCTTCAACTTATGAGCGGTCGACGATTCATGCGGCC 120
DB 61 CGGATGAGAGATACATCATCTTCAACTTATGAGCGGTCGACGATTCATGCGGCC 120

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 31, 2005, 00:02:41 ; Search time 3559 Seconds

(without alignments)
9016.066 Million cell updates/sec

Title: US-10-042-059b-1

Perfect score: 843
Sequence: 1 atgagcttatgagccaga.....tctcctctgctgactag 843

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :
EST:
1: gb_esc1:
2: gb_esc2:
3: gb_esc3:
4: gb_esc4:
5: gb_esc5:
6: gb_esc6:
7: gb_esc7:
8: gb_esc8:
9: gb_esc9:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	299	35.5	1185	9	CNS06TSP
2	176.6	20.9	937	9	CNS06TSP
3	176.6	20.9	1052	9	CNS06TSP
4	140	16.3	948	7	CNS06TSP
5	137	16.3	574	8	CNS06TSP
6	132.8	15.8	738	7	CNS06TSP
7	132.8	15.8	816	6	CNS06TSP
8	125.4	14.9	788	7	CNS06TSP
9	121.6	14.4	660	5	CNS06TSP
10	116.8	13.9	798	6	CNS06TSP
11	116.2	13.8	652	9	CNS06TSP
12	113.8	13.5	747	9	CNS06TSP
13	113.2	13.4	710	7	CNS06TSP
14	111.2	13.2	1222	3	CNS06TSP
15	107.4	12.7	827	3	CNS06TSP
16	106	12.6	988	1	CNS06TSP
17	105.8	12.6	602	6	CNS06TSP
18	105.6	12.5	903	7	CNS06TSP
19	104.6	12.5	527	7	CNS06TSP
20	104.6	12.4	606	4	CNS06TSP
21	103	12.2	867	6	CNS06TSP
22	102.8	12.2	727	6	CNS06TSP
23	102.2	12.1	680	6	CNS06TSP
24	102.2	12.1	680	6	CNS06TSP

25	100.8	12.0	551	5	BO701847
26	99.4	11.8	605	7	CN913779
27	98.6	11.7	607	8	BZ296833
28	98	11.6	495	1	AA786587
29	97.8	11.6	821	7	CK196770
30	97.4	11.6	837	7	CO366854
31	97	11.5	596	8	BZ298767
32	97	11.5	781	7	CF667838
33	96.6	11.5	853	7	CO366784
34	96.2	11.4	657	9	AG273371
35	96	11.4	812	7	CN148636
36	95	11.3	885	7	CV274873
37	94.8	11.2	857	7	CF635400
38	94.6	11.2	480	7	CV002646
39	94.4	11.2	707	6	CA239381
40	94.2	11.2	785	7	CNS23137
41	93.6	11.1	552	4	BT18858
42	93.6	11.1	1265	3	CNS0A6WL
43	93.6	11.1	1329	3	CNS0A727
44	93.2	11.1	698	2	BE577489
45	93.2	11.1	830	6	CB622596

ALIGNMENTS

RESULT 1	CNS06TSP	1185 bp	DNA	linear	GSS 05-UTL-2001
LOCUS	CNS06TSP				
DEFINITION	T7 end of clone AW0A029A11 of library AW0A from strain CL18 89 of				
ACCESSION	AL414131				
VERSION	AL414131.1	GI:12186949			
KEYWORDS					
SOURCE					
ORGANISM	Yarrowia lipolytica				
REFERENCE	Yarrowia lipolytica				
AUTHORS	Soulier, J.L., Aigle, M., Artiguenave, F., Blandin, G., Boivin-Fukuhara, M., Bon, E., Broclet, P., Casaregola, S., de Montigny, J., Dujon, B., Durand, P., Lepingle, A., Lorente, B., Malpertuy, A., Neuvéglise, C., Ozier-Kalogiroupolou, O., Pocher, S., Saurin, M., Tekala, F., Toffano-Nicolas, C., Wesolowski-Louvel, M., Wincker, P., and Weissenbach, J.				
TITLE	Genomic exploration of the hemiascomycetous yeasts: 1. A set of				
JOURNAL	Yeast species for molecular evolution studies				
MEDLINE	FEBS Lett. 487 (1), 3-12 (2000)				
PUBMED	20584711				
REFERENCE	1 (bases 1 to 1185)				
AUTHORS	Casaregola, S., Neuvéglise, C., Lepingle, A., Bon, E., Feynrol, C., Artiguenave, F., Wincker, P., and Galliard, C., Pocher, S., Saurin, M., Tekala, F., Toffano-Nicolas, C., Wesolowski-Louvel, M., Wincker, P., and Weissenbach, J.				
TITLE	Genomic exploration of the hemiascomycetous yeasts: 17. Yarrowia				
JOURNAL	lipolytica				
MEDLINE	FEBS Lett. 487 (1), 95-100 (2000)				
PUBMED	20584727				
REFERENCE	3 (bases 1 to 1185)				
AUTHORS	Genoscope.				
TITLE	Direct Submission				
JOURNAL	Submitted (07-SEP-2000) Genoscope - Centre National de Sequencage,				
COMMENT	2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail: segre@genoscope.cns.fr - Web: www.genoscope.cns.fr)				

This GSS is part of a random genomic sequencing program of thirteen yeast species: *Saccharomyces bayanus* var. *uvatum*, *Saccharomyces exiguus*, *Saccharomyces servazii*, *Zygosaccharomyces rouxii*, *Saccharomyces kluyveri*, *Kluyveromyces thermotolerans*, *Kluyveromyces lactis* var. *lactis*, *Kluyveromyces marxianus* var. *marxianus*, *Pichia angusta*, *Debaryomyces hansenii* var. *hansenii*, *Pichia sorbitophila*, *Candida tropicalis* and *Yarrowia lipolytica*. Genomic inserts of 3 to 5 kb were prepared and both extremities were sequenced. See keywords for description of this sequence and for the sequence of

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2005, 08:01:45 ; Search time 88 Seconds
(without alignment)

1230.601 Million cell updates/sec

Title: US-10-042-059b-2

Sequence: 1 MDPKPEYVLGDIRDALV.....DATQKSGGVDRFLSSGLY 280

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

1: geneseqp19808:*\n2: geneseqp19908:*\n3: geneseqp20008:*\n4: geneseqp20018:*\n5: geneseqp20028:*\n6: geneseqp20038:*\n7: geneseqp20048:*\n8: geneseqp20058:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1434	100.0	280	3	AA011451 H. polym
2	921	64.2	268	8	ADP98857 C. albica
3	587.5	41.0	266	8	ADR86097 Apep9111
4	581.5	40.6	263	6	AA019566 M. sterili
5	508.5	35.5	257	6	ABU08093 Rice chor
6	481.5	33.6	316	5	AB01515 Herbicida
7	481.5	33.6	331	6	ABU08092 Corn chor
8	481.5	33.6	331	6	ADU55262 Corn chor
9	481.5	33.6	334	5	AB01515 Herbicida
10	466	32.5	334	5	ABU08097 A. thalia
11	464	32.4	334	5	ADU55271 Arabidops
12	464	32.4	334	5	ADU55271 Arabidops
13	423	29.5	261	6	ABU08094 Soybean c
14	423	29.5	261	6	ADU55266 Soybean c
15	399	27.8	265	5	AA049654 A. thalia
16	399	27.8	265	5	AB01515 Herbicida
17	399	27.8	265	6	ABU08098 A. thalia
18	399	27.8	265	6	ADU55272 Arabidops
19	399	27.8	265	6	ADU55272 Arabidops
20	391	27.3	319	3	AA060003 Arabidops
21	391	27.3	325	3	AA060002 Arabidops
22	391	27.3	347	3	AA060001 Arabidops
23	313	21.8	207	6	ABU08095 Wheat cho
24	313	21.8	207	6	ADU55268 Wheat cho
25	221	15.4	154	6	ABU08090 Protein e

25	221	15.4	154	8	ADU55258 Wheat cho
27	197.5	11.8	132	6	ABU08089 Protein e
28	197.5	11.8	132	6	ADU55256 Soybean c
29	132	9.2	93	6	ABU08087 Protein e
30	132	9.2	93	6	ADU55252 Corn chor
31	122	8.5	62	6	ABU08088 Protein e
32	122	8.5	62	6	ADU55254 Rice chor
33	114.5	8.0	446	2	AA020059 B. burgdo
34	114.5	8.0	469	2	AA020058 B. burgdo
35	105.5	7.4	1638	8	ADK71824 Human kin
36	103.5	7.2	1732	8	ADP95103 Rat serin
37	103.5	7.2	1732	8	ADP89992 Rat serin
38	103	7.2	724	4	AB015384 Novel hum
39	103	7.2	724	4	AB017531 Novel hum
40	102	7.1	552	6	AB017531 Novel hum
41	100.5	7.0	1664	7	ADK47740 Human NOV
42	100.5	7.0	1664	8	ADJ79010 Human NOV
43	100.5	7.0	1719	5	AA021707 Human NOV
44	100.5	7.0	1732	8	ADP89991 Human NOV
45	100.5	7.0	1732	8	ADP89991 Human NOV

ALIGNMENTS

RESULT 1	AA011451	standard; protein; 280 AA.
XX	AA011451	
AC	AA011451	
XX	12-SEP-2003 (revised)	
DT	01-MAR-2001 (first entry)	
XX	H. polymorpha chorismate mutase protein.	
DB	Chorismate mutase; prephenate; selection marker; auxotrophic yeast.	
XX	Pichia angusta.	
OS	WO200065071-A1.	
PN	02-NOV-2000.	
PD	27-APR-2000; 2000WO-EP003844.	
XX	27-APR-1999; 99DE-01019124.	
PR	(RHEI-) RHEIN BIOTECH GES NEUE BIOTECHNOLOGISCHE.	
XX	Gellissen G, Braus G, Pries R, Krappmann S, Strasser AM,	
XX	WPI: 2000-687355/67.	
DR	N-PSDB; AAC81949.	
XX	Nucleic acids encoding chorismate mutase, useful for preparing an	
PT	auxotrophic selection system for recombinant yeast and for recombinant	
PT	protein expression.	
XX	Claim 1h; Page 57-58; 63pp; German.	
PS	This invention describes novel nucleic acids (I) that encode a	
XX	polypeptide (II) with chorismate mutase (CM) activity (or its	
CC	complementary strand). CM catalyzes conversion of chorismate to	
CC	prephenate, an essential precursor for Phe and Tyr. (I) is a selection	
CC	marker for construction of corresponding auxotrophic yeast (requiring Phe	
CC	and Tyr) that are used for recombinant production of proteins. (I) allows	
CC	selection of transformed yeast on simple media. (Updated on 12-SEP-2003	
CC	to standardise OS field)	
XX	Sequence 280 AA;	
SO	Query Match	100.0%; Score 1434; DB 3; Length 280;

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2005, 10:14:39 ; Search time 29 Seconds

(without alignments)
720.750 Million cell updates/sec

Title: US-10-042-059b-2

Perfect score: 1434
Sequence: 1 MPMKPEYLDIGNIRDALV.....DATKSGGVDRFLSSGLY 280

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgm2_6/prodata/1/1aa/5A_COMB.pep:*
2: /cgm2_6/prodata/1/1aa/5B_COMB.pep:*
3: /cgm2_6/prodata/1/1aa/6A_COMB.pep:*
4: /cgm2_6/prodata/1/1aa/6B_COMB.pep:*
5: /cgm2_6/prodata/1/1aa/6C_COMB.pep:*
6: /cgm2_6/prodata/1/1aa/backfillseq1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the total being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	921	64.2	315	4	US-09-248-796A-18213
2	508.5	35.5	257	4	US-09-454-279-14
3	481.5	33.6	331	4	US-09-454-279-12
4	464	32.4	334	4	US-09-454-279-21
5	423	29.5	261	4	US-09-454-279-16
6	399	27.8	265	4	US-09-454-279-22
7	313	21.8	207	4	US-09-454-279-18
8	221	15.4	154	4	US-09-454-279-8
9	197.5	13.8	132	4	US-09-454-279-6
10	132	9.2	93	4	US-09-454-279-2
11	122	8.5	62	4	US-09-454-279-4
12	94	6.6	391	4	US-09-270-767-42309
13	94	6.6	502	4	US-09-635-872A-13
14	94	6.6	502	4	US-09-636-077A-13
15	94	6.6	502	4	US-09-636-060C-13
16	94	6.6	502	4	US-09-986-552-13
17	94	6.6	502	4	US-09-636-596C-13
18	94	6.6	502	4	US-10-023-894-16
19	94	6.6	502	4	US-10-306-686-13
20	92	6.4	717	4	US-09-583-110-4629
21	92	6.4	721	4	US-09-107-433-2819
22	90	6.3	292	4	US-09-107-532A-3741
23	89.5	6.2	640	4	US-09-248-796A-16129
24	89	6.2	335	4	US-09-934-901-6
25	89	6.2	335	4	US-09-934-868-16
26	89	6.2	335	4	US-10-321-210-6
27	89	6.2	335	4	US-10-320-874-6

28	88	6.1	469	4	US-09-248-796A-14616	Sequence 14616, A
29	87.5	6.1	626	4	US-09-106-872A-4	Sequence 4, App1
30	87	6.1	1871	2	US-08-694-863-1	Sequence 1, App1
31	87	6.1	1871	3	US-09-349-546-1	Sequence 1, App1
32	87	6.1	1871	4	US-09-502-831-1	Sequence 1, App1
33	85	5.9	883	4	US-08-982-430-1	Sequence 1, App1
34	85	5.9	1165	2	US-08-640-389A-11	Sequence 11, App1
35	85	5.9	1165	3	US-09-093-814-1	Sequence 11, App1
36	85	5.9	1165	3	US-08-618-957A-11	Sequence 11, App1
37	85	5.9	1165	4	US-10-095-929-11	Sequence 2, App1
38	85	5.9	1221	4	US-08-982-430-2	Sequence 11, App1
39	84.5	5.9	234	4	US-09-248-796A-19541	Sequence 2782, A
40	84.5	5.9	284	4	US-09-710-279-2782	Sequence 19541, A
41	84.5	5.9	610	4	US-09-710-279-482	Sequence 482, App
42	84.5	5.9	614	3	US-09-134-001C-3061	Sequence 3061, App
43	84.5	5.9	625	4	US-09-710-279-2464	Sequence 2464, App
44	84.5	5.9	631	3	US-09-134-001C-4605	Sequence 4605, App
45	84.5	5.9	661	4	US-09-710-279-274	Sequence 274, App

ALIGNMENTS

RESULT 1
US-09-248-796A-18213
Sequence 18213, Application US/09248796A
Patent No. 6747137
GENERAL INFORMATION:
APPLICANT: Keith Weinstein et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBIC
TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.132
CURRENT APPLICATION NUMBER: US/09/248,796A
CURRENT FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: US 60/074,725
PRIOR FILING DATE: 1998-02-13
PRIOR APPLICATION NUMBER: US 60/096,409
PRIOR FILING DATE: 1998-08-13
NUMBER OF SEQ ID NOS: 28208
SEQ ID NO 18213
LENGTH: 315
TYPE: PRT
ORGANISM: Candida albicans
US-09-248-796A-18213

Query Match 64.2%; Score 921; DB 4; Length 315;
Best Local Similarity 65.6%; Pred. No. 4.7e-88;
Matches 170; Conservative 44; Mismatches 45; Indels 0; Gaps 0;

QY 1 MPMKPEYLDIGNIRDALVNMEDTIIENFIRSOFPYASPSYKNOPIPNFDSFLDW 60
DB 48 MPMKPEYLDIGNIRDALVNMEDTIIENFIRSOFPYASPSYKNOPIPNFDSFLDW 107
QY 61 LLSOHERHSOVRADPDEVPFPNVLEKTELPKINYSVLAASVADENINKEIKIYT 120
DB 108 ALLQLEVASQRRREAPDETPFPDQKTPILPKNYKIAKSDENVSEIMKFFV 167
QY 121 SEIVGIAAGSEODNLSGAMADIECLQSIRRIHFRFYAAEKFISECKIVDLIKK 180
DB 168 DIIIVQVSGGQDQKENSASTCIEQALSRHFKFYAAEKYQSDKPLIYILID 227
QY 181 RVEGIEALITNAEYKRIILRLLEKRAYGDTPLKFTQHIQSKYKPEVIKYIDPVI 240
DB 228 KVKVIGENSITNSAVEQKILRIIYVAESYGVDPISIKFGQNVQSKVPEVIAMLYQDWII 287
QY 241 PLTKKVEVDVYLRLLEDER 259
DB 288 PLTKKVEIDYILRLLEDED 306

RESULT 2
US-09-454-279-14
Sequence 14, Application US/09454279

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2005, 10:17:14 ; Search time 75 Seconds
(without alignments)
1237.947 Million cell updates/sec

Title: US-10-042-059b-2

Perfect score: 1434
Sequence: 1 MDPKPEYVLIDGNIRDALV.....DDATKSGGYVDRFLSSGLY 280

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1413372 seqs, 331592847 residues

Total number of hits satisfying chosen parameters: 1413372

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications AA.*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCF_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCFUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1434	100.0	280	US-10-042-059a-2	Sequence 2, Appl1
2	921	64.2	268	US-10-741-849-7032	Sequence 7032, Ap
3	508.5	35.5	257	US-09-454-279-14	Sequence 14, Appl
4	508.5	35.5	257	US-10-624-061-14	Sequence 14, Appl
5	508.5	35.5	313	US-10-437-963-196439	Sequence 196439, A
6	497.5	34.6	313	US-10-425-114-68586	Sequence 68586, A
7	495.5	34.6	315	US-10-424-599-266662	Sequence 266662, A
8	486	33.6	350	US-10-767-701-45917	Sequence 45917, A
9	481.5	33.6	331	US-09-454-279-12	Sequence 12, Appl
10	481.5	33.6	331	US-10-624-061-12	Sequence 12, Appl
11	464	32.4	334	US-09-454-279-21	Sequence 21, Appl
12	464	32.4	334	US-10-624-061-21	Sequence 21, Appl
13	423	29.5	261	US-09-454-279-16	Sequence 16, Appl

14	423	29.5	261	US-10-624-061-16	Sequence 16, Appl
15	423	29.5	297	US-10-425-114-48703	Sequence 48703, A
16	419	29.2	261	US-10-424-599-235878	Sequence 235878, A
17	408.5	28.5	255	US-10-437-963-151855	Sequence 151855, A
18	399	27.8	265	US-09-454-279-22	Sequence 22, Appl
19	399	27.8	265	US-10-624-061-22	Sequence 22, Appl
20	374.5	26.1	313	US-10-425-114-65608	Sequence 65608, A
21	373	26.0	313	US-10-425-114-65608	Sequence 65608, A
22	356	24.8	306	US-10-437-963-186047	Sequence 186047, A
23	348	24.3	316	US-10-437-963-186040	Sequence 186040, A
24	344.5	24.0	316	US-10-425-114-52677	Sequence 52677, A
25	328	22.9	199	US-10-425-114-53606	Sequence 43606, A
26	313	21.8	207	US-09-454-279-18	Sequence 18, Appl
27	313	21.8	207	US-10-624-061-18	Sequence 18, Appl
28	296.5	20.7	489	US-10-437-963-103185	Sequence 103185, A
29	282	19.7	210	US-10-767-701-32934	Sequence 32934, A
30	221	15.4	154	US-09-454-279-8	Sequence 8, Appl1
31	221	15.4	154	US-10-624-061-8	Sequence 8, Appl1
32	197.5	13.8	132	US-09-454-279-6	Sequence 6, Appl1
33	197.5	13.8	132	US-10-624-061-6	Sequence 6, Appl1
34	178	12.4	109	US-10-424-599-224668	Sequence 224668, A
35	157.5	11.0	97	US-10-425-114-38176	Sequence 38176, A
36	132	9.2	93	US-09-454-279-2	Sequence 2, Appl1
37	132	9.2	93	US-10-624-061-2	Sequence 2, Appl1
38	122	8.5	62	US-09-454-279-4	Sequence 4, Appl1
39	122	8.5	62	US-10-624-061-4	Sequence 4, Appl1
40	106.5	7.4	83	US-10-424-599-272410	Sequence 272410, A
41	104.5	7.3	89	US-10-425-114-52545	Sequence 52545, A
42	102	7.1	552	US-10-136-960-6	Sequence 6, Appl1
43	100.5	7.0	458	US-09-771-161A-137	Sequence 137, Appl
44	100.5	7.0	1664	US-10-210-130-102	Sequence 102, App
45	100.5	7.0	1719	US-10-288-798-2	Sequence 2, Appl1

ALIGNMENTS

RESULT 1
US-10-042-059a-2
Sequence 2, Application US/10042059A
Publication No. US20020197704A1
GENERAL INFORMATION:
APPLICANT: Rhein Biotech Gesellschaft fur neue biotechnologische Prozesse und
APPLICANT: Produktum RMH
TITLE OF INVENTION: Nucleic Acid Molecule Containing a Nucleic Acid Coding for a Pol
FILE REFERENCE: PA30558US-019
CURRENT APPLICATION NUMBER: US/10/042,059A
CURRENT FILING DATE: 2001-10-25
PRIOR APPLICATION NUMBER: DE 199 19 124.7
PRIOR FILING DATE: 1999-04-27
NUMBER OF SEQ ID NOS: 3
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO: 2
LENGTH: 280
TYPE: PRT
ORGANISM: Hansenula polymorpha
US-10-042-059a-2

Query Match	Score	DB 13:	Length	280:
Best Local Similarity	100.0%	Pred. No. 8.7e-120:	Indels	0: Gaps
Matches	280:	Conservative	0:	Mismatches
1	MDPKPEYVLIDGNIRDALVLRMEDTIFNFIERSQFASPSYKVNQPIPNFGSFLDM	60		
1	MDPKPEYVLIDGNIRDALVLRMEDTIFNFIERSQFASPSYKVNQPIPNFGSFLDM	60		
61	LSOHERIRISQVRRYDAPDEVPFNVLEKTFLPKINYPSTLASVADENINKEIKITYT	120		
61	LSOHERIRISQVRRYDAPDEVPFNVLEKTFLPKINYPSTLASVADENINKEIKITYT	120		
121	SEIVPPIAGSGEGEENLASCMAIDIECCQSISRIRHFRPVALEKTSIEGKIYDLIKK	180		

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2005, 10:10:35 ; Search time 26 Seconds
(without alignments)
1036.181 Million cell updates/sec

Title: US-10-042-059b-2

Perfect score: 1434

Sequence: 1 MDPFKPERVLDIGNIRDLV.....DDATQKSGGVDRPLSSGLY 280

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing:

Maximum Match 0%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	776.5	54.1	256	2 A45921	Chorismate mutase
2	558.5	38.9	251	2 T37784	probable chorismat
3	464	32.4	334	2 S38958	chorismate mutase
4	399	27.8	265	2 T50796	chorismate mutase
5	166	11.6	82	2 T14902	chorismate mutase
6	119	8.3	70	2 T14901	chorismate mutase
7	114.5	8.0	469	2 C70109	hypothetical prote
8	104	7.3	313	2 B75138	tetrahydropteroyl
9	104	7.3	336	2 S04739	site-specific DNA-
10	103.5	7.2	1732	2 T14039	protein kinase (EC
11	99	6.9	1033	2 T37715	actin-interacting
12	98	6.8	1702	2 T14050	protein kinase (EC
13	97.5	6.8	1286	2 T37366	RNA polymerase sub
14	97	6.8	323	2 D90527	glycerol-3-phospha
15	97	6.8	1167	2 E96963	DNA polymerase III
16	96.5	6.7	829	2 S72366	DNA topoisomerase
17	96.5	6.7	932	2 P84465	hypothetical prote
18	96	6.7	808	2 H64474	hypothetical prote
19	96	6.7	1286	1 H36845	DNA-directed RNA p
20	96	6.7	1286	2 T28521	DNA-directed RNA p
21	96	6.7	1286	2 A72161	M6R protein - vari
22	95.5	6.7	560	1 F69059	arginine-tRNA liga
23	94.5	6.6	790	2 T19683	hypothetical prote
24	94.5	6.6	1287	1 RNZ447	DNA-directed RNA p
25	94.5	6.6	7829	2 T15789	hypothetical prote
26	94	6.6	311	1 C69952	probable ribosoma
27	94	6.6	502	2 AF1310	probable thermosta
28	94	6.6	612	2 S73611	glucose inhibited
29	93	6.5	307	2 A75020	hypothetical prote

30	92	6.4	992	2 T39315	hypothetical prote
31	91.5	6.4	289	2 D70440	DNA polymerase I 3
32	91.5	6.4	456	2 G81408	probable outer mem
33	91.5	6.4	474	2 S38333	glutathione synth
34	91.5	6.4	663	2 AP1799	hypothetical prote
35	91.5	6.4	913	2 T15474	hypothetical prote
36	91	6.3	345	2 C81300	probable helix-tur
37	91	6.3	377	2 B90437	hypothetical prote
38	91	6.3	398	2 B97341	glycosyltransferas
39	91	6.3	551	2 B90247	DNA repair protein
40	91	6.3	596	2 G82874	conserved hypothet
41	90.5	6.3	475	2 T33979	hypothetical prote
42	90	6.3	431	2 H71172	hypothetical prote
43	90	6.3	445	2 T23199	hypothetical prote
44	89.5	6.2	583	2 C90327	n-methylhydantoina
45	89.5	6.2	859	2 T06352	lipoxigenase (EC 1

ALIGNMENTS

RESULT 1
A45921
chorismate mutase (EC 5.4.99.5) - Yeast (*Saccharomyces cerevisiae*)
N:Alternate names: protein YP9499.15c; protein YPR060c
C:Species: *Saccharomyces cerevisiae*
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 09-Jul-2004
C:Accession: A45921; S54081
R:Schmidheini, T.; Sperisen, P.; Paravicini, G.; Huettler, R.; Braus, G.
J. Bacteriol. 171, 1245-1253, 1989
A:Title: A single point mutation results in a constitutively activated and feedback-free
A:Reference number: A45921; PMID:89155418; PMID:2646272
A:Accession: A45921
A:Molecule type: DNA
A:Residues: 1-256 <SCH>
R:Cross-references: UNIPROT:P32178; GB:M24517; NID:G295576; PIDN:AAB59309.1; PID:G29557
R:Badcock, K.; Churcher, C.M.
submitted to the EMBL Data Library, May 1995
A:Reference number: S54059
A:Accession: S54081
A:Molecule type: DNA
A:Residues: 1-256 <BAD>
R:Cross-references: EMBL:Z49219; NID:G805025; PIDN:CA489177.1; PID:G805040; MIPS:YPR06C
A:Experimental source: strain AB972
C:Genetics:
A:Gene: SGD:ARO7
A:Cross-references: SGD:S0006264; MIPS:YPR060c
A:Map position: 16R
C:Function:
A:Description: intramolecular transferase; isomerase
A:Pathway: aromatic amino acid biosynthesis
C:Superfamily: chorismate mutase of the AroQ class, eukaryotic type
C:Keywords: aromatic amino acid biosynthesis; intramolecular transferase; isomerase

Query Match 54.1%; Score 776.5; DB 2; Length 256;
Best Local Similarity 58.7%; Pred. No. 2.1e-50;
Matches 152; Conservative 40; Mismatches 62; Indels 5; Gaps 2;

QY 1 MDPFKPERVLDIGNIRDLVLRMEDTIIINFIERSQFVSPSYKYNQ--FPINFDGSPFL 58
DB 1 MDPFKPERVLDIGNIRDLVLRMEDTIIINFIERSQFVSPSYKYNQ--FPINFDGSPFL 60
QY 59 DWLISQHERIHSQVRRYAPDEVPFVPLEKTPLPKINYSVTLASVYADEINVKKILKI 118
DB 61 DWLISQHERIHSQVRRYAPDEVPFVPLEKTPLPKINYSVTLASVYADEINVKKILKI 120
QY 119 YVSEITVPGIAGSGEQENIGSCAMADIECIQSRRRIHGFVFAEVRISGDITVDLI 178
DB 121 YIEKTIPLISKRDGDKNNFGSVATRDIECIQSLSRRRIHGFVFAEVRISGDITVDLI 180
QY 179 KRDVEGIEALITNAVEKRIIDRLLEKGRAYGDTPTLKFTHIOSKYKVEYIVKIDYDF 238
DB 181 KSKDVEGIMKNITNSAVBEKILIERLTAKQAYGVDP---NBSGRRITPVELIVKIKYEI 237

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2005, 09:04:29 ; Search time 85 Seconds
(without alignment)

1686.849 Million cell updates/sec

Title: US-10-042-059B-2

Sequence: 1 MDPFKPEVLDLGNIRDALV.....DDATQKSGGYVDRFLSSGLY 280

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot 03:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1434	100.0	280	Q9P4D8	O9P4D8 pichia angu
2	907	63.2	267	Q6BK64	Q6BK64 debaryomyc
3	808	56.3	256	Q6C5J7	Q6C5J7 yarrowia li
4	776.5	54.1	256	1 CHMU YEAST	P34178 saccharomyc
5	752.5	52.5	260	2 Q6CVY3	Q6CVY3 kluyveromyc
6	719	50.1	260	2 Q6FLZ7	Q6FLZ7 candida gla
7	702.5	49.0	259	2 Q75BGS	Q75BGS ashyba goss
8	594	41.4	269	2 Q75BGS	Q75BGS ashyba goss
9	581.5	40.6	263	2 Q6L800	Q6L800 neurospora
10	579	40.4	267	2 Q9Y7B2	Q9Y7B2 emericella
11	558.5	38.9	251	1 CHMU SCHPO	Q11739 schizosacch
12	508.5	35.5	294	2 Q93Y60	Q93Y60 oryza sativ
13	481.5	33.6	316	2 Q9C544	Q9C544 arabidopsis
14	475	33.1	295	2 Q96VZ8	Q96VZ8 cryptococcu
15	466	32.5	334	1 CHMU ARATH	P42738 arabidopsis
16	465	32.4	316	2 Q9STB2	Q9STB2 lycopersico
17	464.5	32.4	355	2 Q9XF60	Q9XF60 arabidopsis
18	408.5	28.5	255	2 Q6H819	Q6H819 oryza sativ
19	399	27.8	265	2 Q9S7H4	Q9S7H4 arabidopsis
20	391	27.3	284	2 Q6Z9B6	Q6Z9B6 oryza sativ
21	290.5	20.3	179	2 Q6F3R1	Q6F3R1 nicotiana t
22	247	17.2	131	2 Q6S861	Q6S861 prunus avu
23	166	11.6	82	2 Q22410	Q22410 petroselinu
24	119	8.3	70	2 Q22409	Q22409 petroselinu
25	114.5	8.0	469	2 Q51102	Q51102 borrelia bu
26	111	7.7	637	2 Q6S1H3	Q6S1H3 bacillus li
27	107	7.5	622	2 Q7NAK6	Q7NAK6 mycoplasma
28	106.5	7.4	413	2 Q8R3P7	Q8R3P7 mus musculu
29	106.5	7.4	1286	2 Q8V4Z4	Q8V4Z4 monkeypox v
30	106	7.4	5289	2 Q88Z03	Q88Z03 lactobacilli
31	105.5	7.4	1638	2 Q81WQ7	Q81WQ7 homo sapien

32	104	7.3	313	2 Q9V086	Q9V086 pyrococcus
33	104	7.3	336	1 MTP2 PROUV	P11409 proteus vul
34	103.5	7.2	1732	2 Q54874	Q54874 rattus norv
35	102.5	7.1	611	2 Q9KX66	Q9KX66 mycoplasma
36	102	7.1	311	2 Q6SH56	Q6SH56 bacillus li
37	102	7.1	345	2 P70904	P70904 borrelia he
38	102	7.1	552	2 Q93EN4	Q93EN4 sarcina ven
39	100.5	7.0	355	2 Q980L9	Q980L9 staphylococ
40	100.5	7.0	357	2 Q8VUY0	Q8VUY0 staphylococ
41	100.5	7.0	1638	2 Q8VX2	Q8VX2 homo sapien
42	100.5	7.0	1719	2 Q86X3	Q86X3 homo sapien
43	99.5	6.9	450	2 Q74K27	Q74K27 lactobacilli
44	99.5	6.9	492	1 K2CO CHICK	Q93532 gallus gall
45	99	6.9	352	2 Q74MT4	Q74MT4 nanorarchaeu

ALIGNMENTS

RESULT 1	ID	Q9P4D8	PRELIMINARY;	PRT;	280 AA.
AC	Q9P4D8				
DT	01-OCT-2000	(TREMBLrel. 15, Created)			
DT	01-OCT-2000	(TREMBLrel. 15, Last sequence update)			
DT	01-MAR-2004	(TREMBLrel. 26, Last annotation update)			
DE	Chorismate mutase (EC 5.4.99.5).				
GN	Name=HAR07;				
OC	Pichia angusta (Yeast) (Hansenula polymorpha).				
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;				
OC	Saccharomycetales; Saccharomycetaceae; Pichia.				
OX	NCBI_TaxID=4905;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=RB11;				
RX	MEDLINE=20353451; PubMed=10894726;				
RX	DOI=10.1128/JB.182.15.4188-4197.2000;				
RA	Krappmann S., Pries R., Gellissen G., Hiller M., Braus G.H.;				
RT	"HAR07 encodes chorismate mutase of the methylotrophic yeast Hansenula				
RT	polymorpha and is derepressed upon methanol utilization.";				
RL	J. Bacteriol. 182:4188-4197(2000).				
DR	EMBL; AF204738; AAF87954.1; -.				
DR	HSSP; P32178; 2CSM.				
DR	GO; GO:0004106; P:chorismate mutase activity; IEA.				
DR	GO; GO:0016853; P:isomerase activity; IEA.				
DR	GO; GO:0009073; P:aromatic amino acid family biosynthesis; IEA.				
DR	InterPro; IPR002701; Chorismate mut.				
DR	InterPro; IPR008238; Chor_mut_AroQ_eu.				
DR	Pfam; PF01817; CM 2; 1.				
DR	PIRSF; PIRSF017318; Chor_mut_AroQ_eu; 1.				
DR	TIGRFAMs; TIGR01802; CM_pl-ycf; 1.				
KW	Isomerase.				
SQ	SEQUENCE 280 AA; 32068 MW; 689D49312CD292D9 CRC64;				
Query Match	100.0%; Score 1434; DB 2; Length 280;				
Best local similarity	100.0%; Pred. No. 3e-92;				
Matches 280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
Qy	1	MDPKPEVLDLGNIRDALVMEETIIENFERSQFYASPSVYKVNQPIPNPFGSFLDW	60		
Db	1	MDPKPEVLDLGNIRDALVMEETIIENFERSQFYASPSVYKVNQPIPNPFGSFLDW	60		
Qy	61	LLSGHERHSQVRDYADPEVPPPNVLEKTFLEPKINYPVSLASVADBINNKSLIKTYT	120		
Db	61	LLSGHERHSQVRDYADPEVPPPNVLEKTFLEPKINYPVSLASVADBINNKSLIKTYT	120		
Qy	61	LLSGHERHSQVRDYADPEVPPPNVLEKTFLEPKINYPVSLASVADBINNKSLIKTYT	120		
Db	61	LLSGHERHSQVRDYADPEVPPPNVLEKTFLEPKINYPVSLASVADBINNKSLIKTYT	120		
Qy	121	SEIVPGIAAGSGREEDNLGSCAMADICLOSIRHIGRVAARAFPSSEDDKIVDLIKK	180		
Db	121	SEIVPGIAAGSGREEDNLGSCAMADICLOSIRHIGRVAARAFPSSEDDKIVDLIKK	180		
Qy	121	SEIVPGIAAGSGREEDNLGSCAMADICLOSIRHIGRVAARAFPSSEDDKIVDLIKK	180		
Db	121	SEIVPGIAAGSGREEDNLGSCAMADICLOSIRHIGRVAARAFPSSEDDKIVDLIKK	180		
Qy	181	RDVGEIALITNAVERKRIIDRLLEKRAVGTDTLTKFTQHIQSKVPEVIVKLYKDFVI	240		
Db	181	RDVGEIALITNAVERKRIIDRLLEKRAVGTDTLTKFTQHIQSKVPEVIVKLYKDFVI	240		
Qy	181	RDVGEIALITNAVERKRIIDRLLEKRAVGTDTLTKFTQHIQSKVPEVIVKLYKDFVI	240		
Db	181	RDVGEIALITNAVERKRIIDRLLEKRAVGTDTLTKFTQHIQSKVPEVIVKLYKDFVI	240		

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 31, 2005, 10:21:29 ; Search time 4655 Seconds
(without alignments)
2914.599 Million cell updates/sec

Title: US-10-042-059B-2
Perfect score: 1434
Sequence: 1 MDMKPEYVLDIGINRDLV.....DDATQSGYVDRFSSGLY 280

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 4708233 seqs, 24227607955 residues
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+g2n:model -DEV=xlh
-O=/cgm2.1/USPTO/seqop/US10042059/runat 30032005_103454_9002/app_query.fasta_1.455
-DB=genmb1 -QMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOEXT=0 -LOOEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -List=45
-DOCLALIGN=200 -THR SCORE=Pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10042059_@CGN_1_1_4200@runat_30032005_103454_9002 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOC
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*

1: gb_da:*
2: gb_hlg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_res:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

No.	Score	Query Match	Length	DB ID	Description
1	1434	100.0	843	6	BD266599 Nucleic a
2	1434	100.0	843	6	AX043758 Sequence
3	1434	100.0	1652	8	AF204738 Pichia an
4	1434	100.0	1655	6	BD266600 Nucleic a

5	1434	100.0	1655	6	AX043760	Sequence
6	921	64.2	948	8	AR548979	Sequence
7	907	63.2	110000	8	CR382138	20
8	808	56.3	110000	8	CR382131	20
9	776.5	54.1	771	8	AT693179	
10	776.5	54.1	2059	8	YSCAR07A	
11	776.5	54.1	43776	8	SC9499X	
12	776.5	54.1	165536	8	SCCHRXVI	
13	752.5	52.5	110000	8	CR382122	07
14	719	50.1	110000	8	CR380957	11
15	702.5	49.0	110000	8	AE016817	01
16	540.5	37.7	869	6	BD178335	
17	540.5	37.7	869	6	AB116236	
18	523	36.5	37301	8	SPAC16E8	
19	511.5	35.7	110000	2	AP006499	05
20	508.5	35.5	1231	6	AR404687	
21	508.5	35.5	1250	8	AK068983	
22	506	35.3	3167	8	AF133241	
23	481.5	33.6	951	8	BT005306	
24	481.5	33.6	1143	8	AK117860	
25	481.5	33.6	1223	6	AR404686	
26	475	33.1	987	8	AF394889	
27	470	32.8	1348	8	AK105512	
28	466	32.5	1005	6	AX506042	
29	465	32.4	1000	8	LA7356	
30	464.5	32.4	1217	6	AR236635	
31	464.5	32.4	1217	6	AR236641	
32	464.5	32.4	1217	6	AX464576	
33	464.5	32.4	1217	6	AX464582	
34	464.5	32.4	1217	6	AF131219	
35	464	32.4	1207	6	AR236633	
36	464	32.4	1207	6	AR236639	
37	464	32.4	1207	6	AX464574	
38	464	32.4	1207	6	AX464580	
39	464	32.4	1207	8	ATCHMUT	
40	451	31.5	1274	8	AY089156	
41	423	29.5	1020	6	AR404688	
42	408.5	28.5	1435	8	AK101220	
43	399	27.8	798	6	AX506277	
44	399	27.8	829	8	AT133840	
45	399	27.8	993	6	AX343935	

ALIGNMENTS

RESULT 1
LOCUS BD266599
DEFINITION Nucleic acid molecule, containing a nucleic acid which codes for a polypeptide with chorismate mutase activity.
ACCESSION BD266599.1 GI:33076367
VERSION BD266599.1
KEYWORDS JP 2002542788-A/1.
SOURCE Pichia angusta
ORGANISM Pichia angusta

REFERENCE 1 (bases 1 to 843)
AUTHORS Gellissen, G., Bries, G., Pries, R., Krapmann, S. and Strasser, A.W.
TITLE Nucleic acid molecule, containing a nucleic acid which codes for a polypeptide with chorismate mutase activity
JOURNAL Patent: JP 2002542788-A 1 17-DEC-2002;
RHEIN BIOTECHE GESELLSCHAFT FUER NEUE BIOTECHNOLOGISCHE PROZESSE UND PRODUKTE MBH

COMMENT
OS Hansenula polymorpha (yeast)
PN JP 2002542788-A/1
PD 17-DEC-2002
PR 27-APR-2000 JP 2000614405
PR 27-APR-1999 DE 199 19 124.7
PT GERD GELLISSEN, GERHARD BRAUS, RALPH PRIES, SVEN KRAPPMANN, PI
ALEXANDER W STRASSER
PC C12N15/09, C07K16/40, C12N1/15, C12N1/19, C12N1/21, C12N5/10 PC
C12N9/90, C12N15/00,

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: March 31, 2005, 10:21:01 ; Search time 584 Seconds
(without alignments)
2638.232 Million cell updates/sec

Title: US-10-042-059b-2
Perfect score: 1434
Sequence: 1 MDMKPEFVLGINIRDALV.....DDATQKSGGVDRFLSSGLY 280

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 4390206 segs, 2959870667 residues
Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame_plus.model -DEV=klh
-O=/cgn2.1/USPTO/spool/US10042053/runat.30032005.103453.8996/app_query.fasta.1.455
-DB=N_Geneseq.16Dec04 -QFMT=fastap -SUFFIX=ing -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsu62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10042053 @CSN_1.1.644 @runat.30032005.103453.8996 -NCP=6 -ICPU=3
-NO WMAP -LARGESUBJECT -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONCLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq.16Dec04:
1: genebegn1808:
2: genebegn1908:
3: genebegn2000s:
4: genebegn2001as:
5: genebegn2002as:
6: genebegn2002bs:
7: genebegn2003as:
8: genebegn2003bs:
9: genebegn2003cs:
10: genebegn2003ds:
11: genebegn2004as:
12: genebegn2004bs:
13: genebegn2004bs:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1434	100.0	843	3	AAC81949
2	1434	100.0	1655	3	AAC81950
3	921	64.2	807	12	ADP98547
4	587.5	41.0	801	13	ADR85510
5	540.5	37.7	869	8	AAL50202

6	517.5	36.1	989	13	ADR84923	Adt84923 Aspergill
7	517.5	36.1	2920	8	ABT17691	Abt17691 Aspergill
8	517.5	36.1	19705	8	ABT19705	Abt19705 Aspergill
9	517.5	36.1	6989	13	ADR84336	Adt84336 Aspergill
10	508.5	35.5	1231	10	ABX93057	Abx93057 Rice chor
11	508.5	35.5	1231	10	ABX93056	Abx93056 Rice chor
12	481.5	33.6	1223	10	ADJ55263	Adj55263 Corn chor
13	481.5	33.6	1223	12	ADJ55261	Adj55261 Corn chor
14	470.5	32.8	658	3	AAI14469	Aai14469 Aspergill
15	466	32.5	1005	6	ABZ12932	Abz12932 Arabidops
16	464.5	32.4	1217	6	ABA91395	Abz1395 Arabidops
17	464.5	32.4	1217	6	ABA91401	Abz1401 Arabidops
18	464	32.4	1207	6	ABA91393	Abz1393 Arabidops
19	464	32.4	1207	6	ABA91399	Abz1399 Arabidops
20	451	31.5	1274	3	AAC33324	Aac33324 Arabidops
21	423	29.5	1020	10	ABX93058	Abx93058 Soybean c
22	423	29.5	1020	12	ADJ55265	Adj55265 Soybean c
23	399	27.8	798	6	ABZ13167	Abz13167 Arabidops
24	399	27.8	993	6	ABA99614	Abz99614 Construct
25	399	27.8	1006	6	ABA99613	Abz99613 A. thalia
26	399	27.8	1006	6	ABA91400	Abz91400 Arabidops
27	399	27.8	1006	6	ABA91394	Abz91394 Arabidops
28	355	24.8	622	13	ACN50132	Acn50132 Cotton no
29	346	24.1	580	13	ACN57510	Acn57510 Cotton gy
30	325	22.7	864	12	ADJ44958	Adj44958 Plant CDN
31	313	21.8	780	10	ABX93059	Abx93059 Wheat cho
32	313	21.8	780	12	ADJ55267	Adj55267 Wheat cho
33	309.5	21.6	5099	6	ABD29643	Abd29643 Arabidops
34	309.5	21.6	5099	6	ABA91396	Abd91396 Arabidops
35	286	19.9	524	13	ADR64235	Adt64235 Cotton CD
36	269	18.8	600	13	ADR64815	Adt64815 Cotton gy
37	245	17.1	542	13	ACN57601	Acn57601 Cotton gy
38	240	16.7	579	12	ADJ55257	Adj55257 Wheat cho
39	240	16.7	579	12	ADJ55257	Adj55257 Wheat cho
40	213.5	14.9	584	13	ACN50205	Acn50205 Cotton no
41	211.5	14.7	525	10	ABX93053	Abx93053 Soybean c
42	211.5	14.7	525	12	ADJ55255	Adj55255 Soybean c
43	209	14.6	258	6	ABT171526	Abt171526 Corn taas
44	195.5	11.6	5176	6	ABA91397	Abz91397 Arabidops
45	191	11.3	507	3	AAI07985	Aai07985 Fusarium

ALIGNMENTS

Result	ID	Accession	Standard	DNA	Length	Description
1	AAC81949	AAC81949	standard	DNA	843 BP.	
XX	XX	XX	XX	XX	XX	
AC	XX	AAC81949				
XX	XX	XX	XX	XX	XX	
DT	XX	15-SEP-2003	(revised)			
DT	XX	01-MAR-2001	(first entry)			
XX	XX	XX	XX	XX	XX	
DE	XX	H. polymorpha	chorismate mutase DNA.			
KM	XX	Chorismate mutase	prephenate; selection marker; auxotrophic yeast; ds.			
OS	XX	Pichia angusta.				
PN	XX	MO200065071-AL.				
XX	XX	XX	XX	XX	XX	
PD	XX	02-NOV-2000.				
XX	XX	XX	XX	XX	XX	
PF	XX	27-APR-2000	2000MO-EP003844.			
XX	XX	XX	XX	XX	XX	
PR	XX	27-APR-1999	99DE-01019124.			
XX	XX	XX	XX	XX	XX	
PA	XX	(RHEI-)	RHEIN BIOTECH GES NEUE BIOTECHNOLOGISCHE.			
PI	XX	Gellissen G,	Braus G, Pries R, Krapmann S, Strasser AW,			
XX	XX	XX	XX	XX	XX	
DR	XX	WPI; 2000-687355/67.				
DR	XX	P-PSDB; AAB11451.				

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: March 31, 2005, 10:42:01 / Search time 195 Seconds
(without alignments)
2349.526 Million cell updates/sec

Title: US-10-042-059B-2

Perfect score: 1434
Sequence: 1 MFMKRETVLDGNIRDALV.....DDATQKSGGVDRFLSSGLY 280

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 1202784 segs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n:models
-O=/cgn2_1/USPTO_SPOOL/US10042055/runat.30032005.103456.9019/app_query.fasta_1.455
-DB=Issued_Patents_NA -QFMT=fastad -SUFFIX=nt1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10042055 @CGN1.1.177 @runat.30032005.103456.9019 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued_Patents_NA:*
1: /cgn2_6/prodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/prodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/prodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/prodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/prodata/1/ina/PCTUS.COMB.seq:*
6: /cgn2_6/prodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	921	66.2	948	4	US-09-248-796A-4110
2	508.5	35.5	1231	4	US-09-454-279-13
3	481.5	33.6	1223	4	US-09-454-279-11
4	464.5	32.4	1217	3	US-09-610-040-4
5	464.5	32.4	1217	3	US-09-610-040-4
6	464.5	32.4	1217	4	US-10-267-763-4
7	464.5	32.4	1217	4	US-10-267-763-4
8	464	32.4	1207	3	US-09-610-040-2
9	464	32.4	1207	3	US-09-610-040-8
10	464	32.4	1207	4	US-10-267-763-2
11	464	32.4	1207	4	US-10-267-763-8
12	423	29.5	1020	4	US-09-454-279-15

13	399	27.8	1006	3	US-09-610-040-3	Sequence 3, Appl1
14	399	27.8	1006	3	US-09-610-040-9	Sequence 9, Appl1
15	399	27.8	1006	4	US-10-267-763-3	Sequence 9, Appl1
16	399	27.8	1006	4	US-10-267-763-9	Sequence 9, Appl1
17	313	21.8	780	4	US-09-454-279-17	Sequence 17, Appl1
18	309.5	21.6	5099	3	US-09-610-040-5	Sequence 5, Appl1
19	309.5	21.6	5099	4	US-10-267-763-5	Sequence 5, Appl1
20	240	16.7	579	4	US-09-454-279-7	Sequence 7, Appl1
21	211.5	14.7	525	4	US-09-454-279-5	Sequence 5, Appl1
22	195.5	13.6	5176	3	US-09-610-040-6	Sequence 6, Appl1
23	195.5	13.6	5176	4	US-10-267-763-6	Sequence 6, Appl1
24	126	8.8	601	4	US-09-454-279-3	Sequence 3, Appl1
25	120.5	8.4	552	4	US-09-454-279-1	Sequence 1, Appl1
26	99	6.9	1664976	4	US-08-916-421B-1	Sequence 1, Appl1
27	99	6.9	1664976	4	US-09-692-570B-1	Sequence 1, Appl1
28	94	6.6	1105	4	US-09-633-872A-17	Sequence 17, Appl1
29	94	6.6	1105	4	US-09-636-077A-17	Sequence 17, Appl1
30	94	6.6	1105	4	US-09-636-060C-17	Sequence 17, Appl1
31	94	6.6	1105	4	US-09-986-552-17	Sequence 17, Appl1
32	94	6.6	1105	4	US-09-636-596C-17	Sequence 17, Appl1
33	94	6.6	1105	4	US-10-023-894-15	Sequence 15, Appl1
34	94	6.6	1105	4	US-10-306-686-17	Sequence 17, Appl1
35	94	6.6	1105	4	US-09-270-767-10717	Sequence 17, Appl1
36	93.5	6.5	59828	4	US-09-949-016-16238	Sequence 16238, A
37	92	6.4	2154	4	US-09-583-110-1968	Sequence 1968, Ap
38	92	6.4	2166	4	US-09-107-433-216	Sequence 216, App
39	90	6.3	879	4	US-09-107-532A-87	Sequence 87, Appl
40	89.5	6.2	1923	4	US-09-248-796A-2026	Sequence 2026, Ap
41	89.5	6.2	2441	3	US-09-051-465-1	Sequence 1, Appl1
42	89.5	6.2	2871	3	US-09-051-465-2	Sequence 2, Appl1
43	89.5	6.2	2871	3	US-09-051-465-4	Sequence 4, Appl1
44	89	6.2	1005	4	US-09-934-901-5	Sequence 5, Appl1
45	89	6.2	1005	4	US-09-934-868-15	Sequence 15, Appl1

ALIGNMENTS

RESULT 1
US-09-248-796A-4110
; Sequence 4110, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBIC
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 4110
; LENGTH: 948
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-4110

Alignment Scores:
Pred. No.: 8.57e-116
Score: 921.00
Percent Similarity: 82.63%
Best Local Similarity: 65.64%
Query Match: 4.23%
DB: 64.23%
Gaps: 0

US-10-042-059B-2 (1-280) x US-09-248-796A-4110 (1-948)
Qy 1 MetasphectylspproglutrhValleuapleuglyaenilleargaspallleuval 20
DB 142 ATGATTATTGAACCAACGAACTGTGCTTATCTTGCAACATCCGTCAGCATTTGCTA 201

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 31, 2005, 11:54:44 ; Search time 622 Seconds
(without alignments)

2724.594 Million cell updates/sec

Title: US-10-042-059B-2

Perfect score: 1434
Sequence: 1 MFMKPEYVLIDGNIRDALY.....DDATKSGGVDFRFLSSGLY 280

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5607317 seqs, 302624599 residues

Total number of hits satisfying chosen parameters: 11214634

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+;p2n.model -DEV=xlh
-O=/cgn2_1/USPTO.spool/US10042059/runat.30032005.103458.9071/app_query.fasta_1.455
-DB=PubMed_Application_NA -QWTF=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPEL=0 -LOOPEXT=0 -UNITS=bites -START=1 -END=-1 -MATRIX=blomsum62
-TRANS=human40.cdi -LIST=45 -DLOCALIGN=200 -THR SCORE=dpct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfco -NORM=ext -HEARSIZE=500 -MILLEN=0
-MAXLEN=200000000 -USRR=US10042059.0CGN_1.1.688.0/runat.30032005.103458.9071
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US10F_NEW_PUB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1434	100.0	843	13	US-10-042-059A-1
2	1434	100.0	1655	13	US-10-042-059A-1
3	921	64.2	807	19	US-10-741-849-6032
4	540.5	37.7	869	18	US-10-472-581-26
5	517.5	36.1	2220	15	US-10-128-714-249
6	517.5	35.1	2987	15	US-10-128-714-249
7	508.5	35.5	1016	18	US-10-437-963-93956
8	508.5	35.5	1231	9	US-09-454-279-13
9	508.5	35.5	1231	17	US-10-624-061-13
10	500.5	34.9	1548	18	US-10-733-930-4455
11	497.5	34.7	1146	17	US-10-425-114-7285
12	495.5	34.6	1423	17	US-10-424-599-123820
13	489.5	34.1	1376	18	US-10-425-115-102347
14	486	33.9	1500	18	US-10-767-701-14353
15	481.5	33.6	1223	9	US-09-454-279-11
16	481.5	33.6	1223	17	US-10-624-061-11
17	479.5	33.4	1430	18	US-10-733-930-4454
18	478.5	33.4	1397	18	US-10-425-115-102346
19	470.5	32.8	658	18	US-10-653-047-6992
20	469.5	32.7	1368	18	US-10-425-115-102349
21	466	32.5	1005	9	US-09-938-842A-737
22	466	32.5	1005	11	US-09-938-842A-737
23	464.5	32.4	1217	14	US-10-267-763-4
24	464.5	32.4	1217	14	US-10-267-763-10
25	464	32.4	1207	14	US-10-267-763-2
26	464	32.4	1207	14	US-10-267-763-8
27	451	31.5	1299	18	US-10-733-930-621
28	423	29.5	1020	9	US-09-454-279-15
29	423	29.5	1020	17	US-10-624-061-15
30	423	29.5	1079	17	US-10-425-114-7639
31	419	29.2	1653	17	US-10-424-599-93036
32	408.5	28.5	768	18	US-10-437-963-93372
33	399	27.8	798	9	US-09-938-842A-972
34	399	27.8	798	11	US-09-938-842A-972
35	399	27.8	1006	14	US-10-267-763-3
36	399	27.8	1006	14	US-10-267-763-9
37	385	26.8	1314	18	US-10-425-115-102351
38	381	26.6	1014	18	US-10-425-115-151470
39	374.5	26.1	1066	17	US-10-425-114-1332
40	374.5	26.1	1083	18	US-10-425-115-115559
41	373	26.0	1197	17	US-10-425-114-27576
42	356	24.8	922	18	US-10-437-963-83564
43	355	24.8	622	18	US-10-021-323-4913
44	351.5	24.5	2154	18	US-10-437-963-83557
45	346	24.1	580	18	US-10-021-323-12291

ALIGNMENTS

RESULT 1
US-10-042-059A-1
Sequence 1, Application US/10042059A
Publication No. US20020197704A1
GENERAL INFORMATION:
APPLICANT: Rhein Botzsch Gesellschaft fur neue biotechnologische Prozesse und
APPLICANT: Prokure mbH
TITLE OF INVENTION: Nucleic Acid Molecule Containing a Nucleic Acid Coding for a Pol
FILE REFERENCE: PA30558US-019
CURRENT FILING DATE: 2001-10-25
PRIOR APPLICATION NUMBER: DE 199 19 124.7
PRIOR FILING DATE: 1999-04-27
NUMBER OF SEQ ID NOS: 3
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 843
TYPE: DNA
ORGANISM: Haemulula polymorpha

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 31, 2005, 10:22:54 ; Search time 7476 Seconds
(without alignments)
1425.628 Million cell updates/sec

Title: US-10-042-059B-2

Perfect score: 1434

Sequence: 1 MDPKMPERVLDLGNTRDALV.....DDATKGGGYDRFLSSGLY 280

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus_p2n.model -DEV=xlh
-Q=/cgn2.1/USPFO/spool/US10042059/runat_30032005_103455_9011/app_query.fasta_1.455
-DB=BST -QFMT=fastap -SUFFIX=trst -MINMATCH=0.1 -LOOPEXT=0 -LIST=45
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCLIGN=200 -THR SCORE=Pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10042055@cgn_1.1.4352@runat_30032005_103455_9011 -NCPU=6 -ICPU=3
-NO MAP -LARGEBUENY -NEG SCORES=0 -WAIT -DSPBLLOC=100 -LONGLOC
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:
1: gb_est1:
2: gb_est2:
3: gb_hic:
4: gb_est3:
5: gb_est4:
6: gb_est5:
7: gb_est6:
8: gb_g881:
9: gb_g882:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	758	52.9	1185	CNS06TSP	AL414111 T7 end of
2	570	39.7	948	CO005472	CO005472 EST793807
3	506.5	35.3	574	BZ303243	BZ303243 KD2284.p1
4	481.5	33.6	1222	AY103806	AY103806 Zea mays
5	472	32.9	788	CO165801	CO165801 FLD1_57_E
6	470	32.8	1329	CNS04727	BX822129 Arabidops
7	469	32.7	937	CNS06TV7	AL415049 T3 end of
8	469	32.7	1052	CNS06XUJ	AL419633 T3 end of
9	465	32.4	1265	CNS0A6WL	BX824460 Arabidops

10	460.5	32.1	710	CF475712	CF475712 RTW2_11
11	458.5	32.0	828	CO120964	CO120964 GR_EB024
12	448	31.2	737	CO005473	CO005473 EST793808
13	442.5	30.9	867	CB619882	CB619882 OSTIB04L
14	436.5	30.4	652	CD423256	CD423256 SAI_28_B1
15	436.5	30.4	660	BX253798	BX253798 BX253798
16	434.5	30.3	821	CK196770	CK196770 FGAS00523
17	434	30.3	738	CF877247	CF877247 trlc075xb
18	434	30.3	815	CB905545	CB905545 trlc075xb
19	418	29.1	657	AG273371	AG273371 Cyanidios
20	416.5	29.0	680	AJ612241	AJ612241 AJ612241
21	415.5	28.0	738	BQ989273	BQ989273 QGF17D13
22	413.5	28.8	785	CNS23137	CNS23137 QGF15M01
23	409	28.5	680	CA080410	CA080410 SCVPM105
24	407	28.4	798	CB619881	CB619881 OSTIB04L
25	404	28.2	988	AJ833468	AJ833468 AJ833468
26	403.5	28.1	676	CD423358	CD423358 SAI_28_B1
27	401.5	28.0	888	AY389761	AY389761 Hyacinthu
28	400.5	27.9	711	BF430733	BF430733 OG04A12T3
29	398.5	27.8	903	CN149106	CN149106 MOUNDI_60
30	394.5	27.5	853	CO365784	CO365784 RTK1_35_C
31	392	27.3	840	CO071620	CO071620 GR_Ea30A
32	392	27.3	865	CO071737	CO071737 GR_Ea30E
33	387.5	27.0	656	CA293493	CA293493 SCSEGLV100
34	387	27.0	866	CO071299	CO071299 GR_Ea29C
35	384.5	26.8	830	CB622596	CB622596 OSTIB091
36	383.5	26.7	795	CF444592	CF444592 EST680937
37	383.5	26.7	857	CF635400	CF635400 zmrw00_0
38	382	26.6	602	CD982471	CD982471 QAK504.x
39	381	26.6	527	CO165812	CO165812 FLD1_57_F
40	380.5	26.5	810	CD427426	CD427426 SAI_30_H0
41	379.5	26.5	885	CV274873	CV274873 MS0174_B2
42	378	26.4	727	CA121888	CA121888 SCIFR105
43	376.5	26.3	698	BE577489	BE577489 L48-2335T
44	374	26.1	526	BE494722	BE494722 WHE1257_A
45	374	26.1	863	CO094261	CO094261 GR_Ea16E

ALIGNMENTS

RESULT 1	CNS06TSP	1185 bp	DNA	linear	GSS 05-JUN-2001
LOCUS	Yarrowia lipolytica				
DEFINITION	T7 end of clone AM0A029A11 of library AM0A from strain CLIB 89 of Yarrowia lipolytica, genomic survey sequence.				
ACCESSION	AL414131				
VERSION	AL414131.1	GI:12186949			
KEYWORDS	GSS.				
SOURCE	Yarrowia lipolytica				
ORGANISM	Yarrowia lipolytica				
REFERENCE	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Dipodascaceae; Yarrowia.				
AUTHORS	1 (bases 1 to 1185) Souciet, J.-L., Aigle, M., Artiguenave, F., Blandin, G., de Montigny, J., Dujon, B., Durrens, P., Lepingle, A., Lorente, B., Maupertuy, A., Neuvéglise, C., Ozier-Ladoire, O., Potier, S., Saurin, W., Tekala, F., Toffano-Nioche, C., Wesolowski-Louvel, M., Wincker, P., and Weissenbach, J.				
TITLE	Genomic exploration of the hemiascomycetous yeasts: 1. A set of yeast species for molecular evolution studies				
JOURNAL	FEBS Lett. 487 (1), 3-12 (2000)				
MEDLINE	20584711				
PUBMED	11152876				
REFERENCE	2 (bases 1 to 1185) Casaregola, S., Neuvéglise, C., Lepingle, A., Bon, E., Feynereol, C., Artiguenave, F., Wincker, P., and Galliardin, C.				
AUTHORS	Genomic exploration of the hemiascomycetous yeasts: 17. Yarrowia lipolytica				
TITLE	FEBS Lett. 487 (1), 95-100 (2000)				
JOURNAL	20584727				
MEDLINE	11152892				
PUBMED	3 (bases 1 to 1185)				
REFERENCE					